



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

A Calculus of Chemical Systems

Citation for published version:

Plotkin, G 2013, A Calculus of Chemical Systems. in V Tannen, L Wong, L Libkin, W Fan, W-C Tan & M Fourman (eds), *In Search of Elegance in the Theory and Practice of Computation: Essays Dedicated to Peter Buneman*. Lecture Notes in Computer Science, vol. 8000, Springer-Verlag GmbH, pp. 445-465.
https://doi.org/10.1007/978-3-642-41660-6_24

Digital Object Identifier (DOI):

[10.1007/978-3-642-41660-6_24](https://doi.org/10.1007/978-3-642-41660-6_24)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

In Search of Elegance in the Theory and Practice of Computation

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



A CALCULUS OF CHEMICAL SYSTEMS

GORDON D. PLOTKIN

1. INTRODUCTION

In recent years various calculi have been proposed for modelling biological systems, typically intracellular pathways. These calculi generally fall into one of two camps: ones based on process calculi, such as Milner’s pi-calculus [24], and rule-based ones. Examples of the former include [31, 32, 30]; examples of the latter include BIOCHAM, κ , BioNet-Gen, and Dynamical Grammars [2, 6, 15, 26]. One positive feature of the rule-based approach is that rules correspond naturally to biological events, with a main example being chemical reactions. However rule-based models usually consist of sets of rules, and these are completely unstructured. In sharp contrast, process calculi have natural means for specifying modularity such as a parallel combinator, to combine subsystems, and the ability to name and parameterise subsystems. They thereby can avoid much redundancy.

Here we aim at combining the naturalness of rules with the modularity of process calculi. We keep both sides of the endeavour simple. Our rules are simple reactions between species; our modularity provides parallelism and the ability to define modules (without parameters). We call the resulting formalism the *Calculus of Chemical Systems* and consciously model it on Milner’s CCS [23], his *Calculus of Communicating Systems*. There are two versions of our calculus, a *qualitative* one and, more importantly, a *quantitative* one where the rules are equipped with rates. In both cases modules are non-recursive; this is more suited to biological systems than the recursive modules of Milner’s CCS.

There are strong arguments in favour of the existence of modules at the cellular level [13, 1], although it is worth remarking that the extent to which biological systems are modular is controversial because of the high degree of interaction between supposed subsystems. Nonetheless, biologists do not describe entire cells but rather particular pathways and they further note recurring patterns; see, for example, [10] on the MAPK cascade. There is also speculation that modules have evolutionary significance and provide robustness to environmental perturbations [13, 1]. In terms of Petri nets, where there has also been much research on modularity, e.g., [19, 11, 28], we combine nets sharing commonly named places (species) but not transitions (reactions). Whereas we provide means for describing modules one can also try to find modules within systems with given unstructured descriptions, often graphical; see [18] for a survey of work along these lines.

Following the CCS paradigm, we look at a number of equivalences, here naturally given via various compositional denotational semantics. In giving these semantics we take up a suggestion of Pedersen [29] to incorporate the various semantics within a

general framework, here using commutative monoids. (This is strongly connected to the use of rings of operators by Mjolsness and Yosiphon [26].)

For the quantitative case, we consider two equivalencies, one based on ordinary differential equations (ODEs), and the other on stochastic matrices. For the qualitative case we consider a transition relation semantics and a Petri net one. As is well known, reactions can also be naturally represented by Petri nets, particularly P/T nets, and, indeed there is little difference between P/T nets and sets of reactions. See [4, 14, 25] for surveys of the now extensive applications of Petri nets to pathway simulation; see [35, 16] for an early chemical reaction formalism using a bipartite graph structure equivalent to P/T nets; and see [8] for a survey of the application of graphical ideas to chemical reaction systems, including two bipartite graph formalisms [17, 5].

We provide a complete axiomatisation of each semantics that we consider. We also consider definability, the main result here is a characterisation of the collection of definable sets of ODEs due to Hárs and Tóth [12, 7]. Continuing to consider the analogy with Milner’s CCS, one could also seek analogues of Hennesy-Milner logic, but we do not do so here.

Since our calculus and P/T nets are ways of specifying sets of reactions, there is a tight relationship between them: the kind of P/T nets we consider form a complete model of the basic axiom system of the calculus, and all such nets are definable, up to isomorphism. This gives the calculus a reading as a modular way to describe Petri nets; conversely, the possibility arises of inputting programs via a graphical input system.

As is already well-known, one can give P/T nets themselves qualitative or quantitative semantics (for the quantitative semantics [14, 36] one has, of course, to equip the transitions with rates). We show that each such direct semantics agrees with the corresponding calculus semantics, by which we mean that the latter factors through the former via the P/T net semantics of the calculus. As the various semantics of the P/T nets are standard, this acts as a verification of the corresponding semantics of the calculus. Another consequence is that properties of the semantics of P/T nets correspond with properties of the semantics of processes of the calculus; therefore any means of establishing that one holds provides a means of establishing that the other does.

The calculus of chemical systems is intended as a simulation language. So, in the cases of the quantitative semantics, one could argue that one is interested in the simulations that processes induce rather than the ODEs, or stochastic matrices, themselves. These simulations are the solutions to the ODEs, or (samples of) the Markov processes induced by the stochastic matrices, and are functionally dependent on given initial states. Unfortunately, such simulations do not have a *prima facie* compositional (here, commutative monoid) structure. Following a standard move in the programming language literature, see, e.g., [37, 9], one could say that the correct equivalence should be an analogue of behavioural (a.k.a. observational, or, more neutrally, contextual) equivalence, that is, having the same behaviour, here simulations, in all contexts. Pleasingly, it turns out that we have full abstraction results in both cases, that is, the denotational and contextual equivalences agree.

Our calculus is intended as a simplest possible rule-based calculus. It should be possible to usefully add modularity to any rule-based calculus. Indeed the grammatical approach in [26] provides what amounts to a parameterised module mechanism. One

can enrich our calculus with facilities for modifications and simple complexes along the general lines of BIOCHAM, and with a construction for (static) compartments: all of these are omnipresent in cellular networks. One then has a calculus briefly described in [29], and called there the *Calculus of Biochemical Systems*. That paper provided further linguistic facilities, for example parameterised modules, several kinds of declarations, and functional reaction rates; the aim was to provide a full-scale modular modelling language. In principle, such further developments can be carried out for any rule-based system.

After some technical preliminaries in Section 2, we present the calculus of chemical systems in Section 3. We begin Section 4 with the commutative monoid semantical framework. We give two semantics for the qualitative version of the calculus, a transition relation one and a P/T net one, and we give three semantics for the quantitative calculus, one using a kind of P/T net equipped with reaction rates, and the differential and stochastic ones, all as mentioned above. Completeness results are given for all of these (Theorems 4.1, 4.2, 4.5, and 4.9, and discussions). Characterisations of definability are given for the (immediately obvious) Petri net case and for the ODE case (Theorem 4.4); the other cases present interesting open problems. Finally the full abstraction result for the ODE case is given by Theorem 4.8, and that for the stochastic case by Theorem 4.12.

2. TECHNICAL PRELIMINARIES

For any set A we write $\mathcal{M}_f(A)$ for the collection of finite multisets on A . For any such multiset X , the multiplicity of $a \in A$ is written $X(a)$ and we write $\|X\|$ for its size, i.e., $\sum_{a \in A} X(a)$. The empty multiset is written as \emptyset ; we may confuse an $a \in A$ with the multiset with unique element a occurring with multiplicity 1; and for any multisets X and Y we write mX and $X + Y$ for the multisets such that $(mX)(a) = mX(a)$ and $(X + Y)(a) = X(a) + Y(a)$, respectively. So, in particular, for natural numbers m_1, \dots, m_r , and $a_1, \dots, a_r \in A$ we have the multiset $m_1a_1 + \dots + m_ra_r$. Finally, for any function $f : A \rightarrow B$, and $a \in A$ and $b \in B$, $f[a \mapsto b]$ denotes the function everywhere equal to f except, possibly, at a where it has the value b .

3. THE CALCULUS

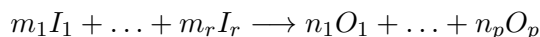
We assume given a nonempty finite set Spec of *species*, ranged over by S, I, O . The set of species names is a parameter to our formalism and may be varied according to the area of application.

The set of *rules*, ranged over by R , is given by:

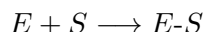
$$R ::= X \rightarrow Y$$

where X and Y are finite multisets of species.

We can use rules to describe chemical reactions, as in:



where the I_i are the reactants and the O_j are the products and the stoichiometry is given by the m_i and the n_j . Rules can also be used to describe complex formation, as in:



where E - S is an assumed name for an E , S -complex, and to describe transport, as in:

$$S \rightarrow \text{nucleus}[S]$$

where $\text{nucleus}[S]$ is an assumed name for S inside the cell nucleus.

The set of (*qualitative*) *processes*, ranged over by P and Q , is then given by the following abstract syntax:

$$P ::= R \mid P|Q \mid \text{NIL} \mid A = P; Q \mid A$$

where A ranges over a countably infinite set ProcId of *process identifiers*. The first case is that of a rule, and the second case $P|Q$ is that of two systems P and Q acting in parallel: as a species name may occur in both they may interact. The third case is the empty system and the fourth case is a local definition: of A as P in Q . The last case is whatever system the identifier A denotes: as such it will normally be a fragment of a larger process in which it receives a definition.

Local definition is a binding construct with A having scope Q in $A = P; Q$. We have the usual notions of the free process identifiers $\text{FP}(P)$ occurring in a process P , of α -equivalence of processes, $P \equiv_\alpha Q$, and of the capture-avoiding substitution $P[Q/A]$ of a process Q for a process identifier A in a process P . We adopt a standard convention and regard α -equivalent processes as identical.

For the set of *quantitative* processes, rules instead have the form

$$X \xrightarrow{r} Y$$

where r is a positive real (i.e., $r > 0$), which gives the rate at which the rule occurs.

There are natural abbreviations. For reversible reactions, we could write

$$X \xleftrightarrow[s]{r} Y$$

as an abbreviation for

$$X \xrightarrow{r} Y \mid Y \xrightarrow{s} X$$

in the quantitative language (and here, and below, we assume the evident corresponding abbreviations for the qualitative language).

For an enzymatic reaction in which an enzyme E converts a substrate S to a product P we could write

$$E : S \xrightarrow{r,s,t} P$$

as an abbreviation for

$$S + E \xleftrightarrow[s]{r} E\text{-}S \mid E\text{-}S \xrightarrow{t} E + P$$

Finally for a reversible enzymatic reaction with different enzymes E and E' in the forwards and backwards directions we could write

$$E : S \xleftrightarrow[r',s',t']{r,s,t} P : E'$$

for

$$E : S \xrightarrow{r,s,t} P \mid E' : P \xrightarrow{r',s',t'} S$$

There is an interesting empirical adequacy question here concerning the design of abbreviations: one needs to be able to use all the abbreviations commonly found in the literature. More ambitiously one may try to generate useful such abbreviations

systematically. For example one may add those suggested in [38] in order to obtain natural, succinct descriptions of a variety of enzyme kinetics.

There is an evident, very simple, equational theory for processes (more precisely, one for each of the qualitative or quantitative cases). It has the standard rules of reflexivity, transitivity and symmetry and substitutivity for the parallel construct, together with the following congruence rule for local definitions:

$$\frac{P = P' \quad Q = Q'}{(A = P; Q) = (A = P'; Q')}$$

The axioms are that the parallel construct and NIL form a commutative monoid:

$$(P_1 \mid P_2) \mid P_3 = P_1 \mid (P_2 \mid P_3) \quad P_1 \mid P_2 = P_2 \mid P_1 \quad P \mid \text{NIL} = P$$

together with:

$$(A = P; Q) = Q[P/A]$$

Every process can be put in the following *canonical form*:

$$A_1 \mid \dots \mid A_m \mid R_1 \mid \dots \mid R_n$$

associating parentheses to the left. We remark that the rule of substitutivity for local definitions is redundant as processes can be put in canonical form without using it.

It follows from Theorem 4.1, below (and from its analogue for the quantitative case) that two such canonical forms are provably equal if, and only if, the same process names and reactions occur in each, possibly in different orders; in other words if the two multisets of process names and reactions obtained from each canonical form are equal. With that we know that processes have *normal forms*, i.e., canonical forms unique up to reordering. The same situation obtains with respect to all the semantics considered below: there are evident canonical forms, and the relevant completeness theorems ensure that they are, in fact, normal forms, i.e., they are unique up to reordering.

4. SEMANTICS

We give a variety of (denotational) semantics for both the qualitative and quantitative calculi. It proves convenient to organise all these under one scheme. Suppose that we have a structure

$$(M, \mid_M, \text{NIL}_M)$$

where \mid_M is a binary operation on M and NIL_M is an element of M , and for every rule R we are given an element of M , its *semantics*, written $M[R]$, slightly abusing notation. A *process environment* is a map

$$\rho : \text{ProcId} \rightarrow M$$

For every process P we give its semantics $M[P](\rho)$ relative to a process environment ρ , by:

$$\begin{aligned} M[R](\rho) &= M[R] \\ M[P \mid Q](\rho) &= M[P](\rho) \mid_M M[Q](\rho) \\ M[\text{NIL}](\rho) &= \text{NIL}_M \\ M[A = P; Q](\rho) &= M[Q](\rho[A \mapsto M[P](\rho)]) \\ M[A](\rho) &= \rho(A) \end{aligned}$$

Note that $M\llbracket P \rrbracket(\rho)$ depends only on the values that ρ assigns to the free process variables of M ; we may therefore write $M\llbracket P \rrbracket$ instead of $M\llbracket P \rrbracket(\rho)$ when M is closed; elements of M of the form $M\llbracket P \rrbracket$ are said to be *definable*. As is usual, denotational semantics commutes with substitution, by which is meant that:

$$M\llbracket P[Q/A] \rrbracket(\rho) = M\llbracket P \rrbracket(\rho[A \mapsto M\llbracket Q \rrbracket(\rho)])$$

always holds.

If M is a commutative monoid then we obtain a model of the above equational theory, meaning that if $P = Q$ is provable then $M\llbracket P \rrbracket(\rho) = M\llbracket Q \rrbracket(\rho)$, for all ρ . Conversely, we say that such a model is *complete* if whenever $M\llbracket P \rrbracket(\rho) = M\llbracket Q \rrbracket(\rho)$ for all ρ then $P = Q$ is provable.

One may instead have a structure $(M, |_M, \text{NIL}_M)$, which is not a commutative monoid, but can be equipped with a congruence \sim such that M/\sim is a commutative monoid: this occurs, for example, in the case of Petri nets. One then has a model in the weaker sense, that if $P = Q$ is provable then $M\llbracket P \rrbracket(\rho) \sim M\llbracket Q \rrbracket(\rho)$ for all ρ , and one defines completeness accordingly.

4.1. Petri Net Semantics. We begin with the qualitative case. Following the above framework we need to give a suitable set Net of Petri nets. We take these to be structures

$$(T, \text{pre}, \text{post})$$

where T , the set of transitions, is a finite subset of $\{0, 1\}^*$ and $\text{pre}, \text{post} : T \rightarrow \mathcal{M}_f(\text{Spec})$. We may write $\text{pre}(t)$, $\text{post}(t)$ as, respectively, $\cdot t$, $t \cdot$. We do not specify a set of places as we use the same set of places, Spec, for all our nets. Note too that places are identified with (basic) species, as is normally done when using Petri nets to describe biochemical processes, and which corresponds to the convention in pathway graphical formalisms of having each species occur just once in the graph. Further, transitions are kept anonymous: this could, of course be changed, but they are normally kept so, and we know no reason to depart from that practice. Our nets are P/T nets with the small differences that we have a fixed set of places and that we do not have an initial marking.

The semantics of rules is given by:

$$\text{Net}\llbracket X \longrightarrow Y \rrbracket = (\{\varepsilon\}, \{\varepsilon \mapsto X\}, \{\varepsilon \mapsto Y\})$$

Next, the function $|_{\text{Net}} : \text{Net}^2 \rightarrow \text{Net}$ is defined by:

$$(T, \text{pre}, \text{post}) |_{\text{Net}} (T', \text{pre}', \text{post}') = (0T \cup 1T', \text{pre}'', \text{post}'')$$

where

$$\text{pre}''(u) =_{\text{def}} \begin{cases} \text{pre}(w) & (\text{if } u = 0w) \\ \text{pre}'(w) & (\text{if } u = 1w) \end{cases}$$

and

$$\text{post}''(u) =_{\text{def}} \begin{cases} \text{post}(w) & (\text{if } u = 0w) \\ \text{post}'(w) & (\text{if } u = 1w) \end{cases}$$

And we define:

$$\text{NIL}_{\text{Net}} = (\emptyset, \emptyset, \emptyset)$$

meaning the constantly \emptyset function in the second and third components.

With these definitions $(\text{Net}, |_{\text{Net}}, \text{NIL}_{\text{Net}})$ is not a commutative monoid, because of slight differences of transitions. However we can define a congruence, putting

$$(T, \text{pre}, \text{post}) \sim (T', \text{pre}', \text{post}')$$

if, and only if, the two nets are isomorphic, meaning that there is a bijection $\theta : T \cong T'$ such that, for all $t \in T$ we have $\text{pre}(t) = \text{pre}'(\theta(t))$ and $\text{post}(t) = \text{post}'(\theta(t))$, and, dividing out by this congruence, we do obtain a commutative monoid, as required.

In a way, the calculus of chemical systems is little more than a way of writing down Petri nets, and this results in some natural formal relations between the two. Concerning definability, as is evident, every net in Net is definable by a program, up to isomorphism. This means that for every such net N there is a closed process P such that $\llbracket P \rrbracket \sim N$. This is the formal correlate of the pragmatic possibility of inputting processes by drawing nets, just as the semantics is the formal correlate of the pragmatic possibility of defining nets by processes.

We also have that the Petri net semantics is complete, and that the canonical forms are unique up to reordering (we then say that they are *normal* forms):

Theorem 4.1. *The following are equivalent for any two processes P and Q :*

- (1) *P and Q can be proven equal.*
- (2) *Any two canonical forms of P and Q are the same up to reordering.*
- (3) *$\text{Net}\llbracket P \rrbracket(\rho) \sim \text{Net}\llbracket Q \rrbracket(\rho)$ for all ρ .*

Proof. Clearly (2) \Rightarrow (1) \Rightarrow (3). So assume (3) and consider two processes P and Q with canonical forms $A_1 \mid \dots \mid A_m \mid R_1 \mid \dots \mid R_n$ and $A'_1 \mid \dots \mid A'_{m'} \mid R'_1 \mid \dots \mid R'_{n'}$. Recalling that Spec is nonempty, choose $S \in \text{Spec}$, and let B_1, \dots, B_n be a list of the process variables occurring freely in either one of P or Q , and let s be the maximum of the number of occurrences of S in an R_j or a $R'_{j'}$. Now choose ρ so that $\rho(B_k)$ is a net with one transition ε such that $\varepsilon(S) = s + k$ and $\varepsilon^-(S) = 0$. Then, as $\text{Net}\llbracket P \rrbracket(\rho) \sim \text{Net}\llbracket Q \rrbracket(\rho)$, one has that the two chosen canonical forms of P and Q are identical up to reordering (we omit further details). \square

There is a variant of the Petri net semantics worth remarking. One only allows nets not containing any transition w with ' $w = w$ ' and takes the semantics of $X \rightarrow X$ to be NIL_{Net} . Everything then goes through as before except that in the proof system one adds the axiom

$$X \rightarrow X = \text{NIL}$$

and in canonical forms one forbids rules of the form $X \rightarrow X$.

In the quantitative case rules have associated positive reals and we instead take our set of quantitative Petri nets, CNet to be structures

$$(T, \text{pre}, \text{post}, \text{rate})$$

where T , pre and post are as before and $\text{rate} : T \rightarrow \mathbb{R}^+$, where \mathbb{R}^+ is the set of positive reals. (Formally, these nets are stochastic nets but for the same small differences as before: there is a fixed set of places and there is no initial marking. However, our nets have no fixed intended interpretation, and, indeed, will receive both an ODE and a stochastic semantics.)

One then makes the evident changes to the semantics for the qualitative case: we do not spell them out here. The evident analogue of Theorem 4.1 holds, with the analogous proof. One also has the evident analogue of the above variant of the Petri net semantics, now with $X \xrightarrow{r} X = \text{NIL}$ as the additional axiom.

4.2. Qualitative Semantics: Transition Relations. We give a semantics of the qualitative calculus in terms of transition relations on *markings*, here finite multisets of species. The idea is that the marking gives the population of each species and the transitions correspond to the occurrence of a finite multiset of reactions, or other rules. So we set TRel to be the set of relations \rightarrow on $\mathcal{M}_f(\text{Spec})$ satisfying the following closure conditions:

$$X \rightarrow X \quad \frac{X \rightarrow Y \quad X' \rightarrow Y'}{X + X' \rightarrow Y + Y'}$$

The first corresponds to the fact that the multiset of rules firing may be empty; the second corresponds to the fact that the firings of two multisets can be combined into their joint firing. Given any relation \rightarrow on $\mathcal{M}_f(\text{Spec})$ there is a least such relation \rightarrow^* containing it and closed under the rules; we have:

$$\rightarrow^* = \{(Z + \sum_{i=1,n} m_i X_i, Z + \sum_{i=1,n} m_i Y_i) \mid X_i \rightarrow Y_i, m_i \in \mathbb{N} (i = 1, n)\}$$

We say that \rightarrow *generates* \rightarrow^* .

The semantics of rules is given by:

$$\text{TRel}[X \rightarrow Y] = \{(Z + kX, Z + kY) \mid Z \in \mathcal{M}_f(\text{Spec}), k \in \mathbb{N}\} (= \{(X, Y)\}^*)$$

and we define a commutative monoid structure on TRel by:

$$\rightarrow \mid_{\text{TRel}} \rightarrow' = \{(kX + k'X', kY + k'Y') \mid k, k' \in \mathbb{N}, X \rightarrow Y, X' \rightarrow' Y'\} (= (\rightarrow \cup \rightarrow')^*)$$

and:

$$\text{NIL}_{\text{TRel}} = \{(X, X) \mid X \in \mathcal{M}_f(\text{Spec})\} (= \emptyset^*)$$

Note that \mid_{TRel} is *absorptive*, i.e., that, for any transition relation \rightarrow , we have:

$$\rightarrow \mid_{\text{TRel}} \rightarrow = \rightarrow$$

We do not know any characterisation of the relations defined by closed processes, other than the obvious one that they are generated by a finite relation. An obvious necessary condition is that, for any X , the set $\{Y \mid X \rightarrow Y\}$ is finite. However there are closed non-finitely generated relations obeying this condition, e.g., $\{(mS, nS) \mid m \geq 1, m \leq n \leq m^2\}$ where the maximum gap size, $m^2 - m$, grows too quickly. This example also satisfies two other obvious conditions: that both the set $\{Y \mid X \rightarrow Y\}$ and its size are recursive in X .

As well as the axioms above, another three hold in TRel :

$$P \mid P = P$$

$$X \rightarrow X = \text{NIL}$$

$$X \rightarrow Y \mid X' \rightarrow Y' = X \rightarrow Y \mid X' \rightarrow Y' \mid X + X' \rightarrow Y + Y'$$

The first corresponds to the absorptiveness of \mid_{TRel} and the others to the above closure conditions. Every process can be put in a canonical form. This is, as before, a composition of process names and rules but where, in addition, no process name or rule occurs

twice, and where no rule is in the closure of the others (considered as forming a relation).

As we now see, with these additional axioms we obtain completeness for the transition relation semantics:

Theorem 4.2. *The following are equivalent for any two processes P and Q (assuming there is an element of Spec not occurring in P or Q):*

- (1) P and Q can be proven equal using the above axioms for the qualitative semantics.
- (2) Any two canonical forms of P and Q are the same up to reordering.
- (3) $\text{TRel}\llbracket P \rrbracket(\rho) = \text{TRel}\llbracket Q \rrbracket(\rho)$ for all ρ .

Proof. As in the case of Theorem 4.1 we need only assume (3) and prove (2). So let $A_1 \mid \dots \mid A_m \mid R_1 \mid \dots \mid R_n$ and $A'_1 \mid \dots \mid A'_{m'} \mid R'_1 \mid \dots \mid R'_{n'}$ be canonical forms of P and Q . For any A_i , first choose S not occurring in P or Q and then choose ρ such that $\rho(A_i) = \{(S, 0)\}^*$ and, otherwise, is NIL_{TRel} . Then $\text{TRel}\llbracket P \rrbracket(\rho)$ is the least transition relation containing (S, \emptyset) and the R_i (considered as ordered pairs of multisets). So, by assumption, (S, \emptyset) also occurs in $\text{TRel}\llbracket Q \rrbracket(\rho)$ and so, as S occurs in no $R'_{j'}$, A_i must be one of the $A'_{j'}$. With this, and the symmetric argument, we see that the A_i and the $A'_{j'}$ are the same.

Next, taking ρ constantly NIL_{TRel} , we see that the R_j and the $R'_{j'}$ generate the same transition relations using the above closure conditions as rules. Observe that if (X, Y) is generated from a set of pairs of multisets, then either $X = Y = \emptyset$ or (X, Y) is in the set or else it is generated from (X', Y') in the set with $\|X'\| + \|Y'\| < \|X\| + \|Y\|$. So as R_1 (say) is not $\emptyset \rightarrow \emptyset$ either it is some $R'_{j'}$, or else it is generated from $R'_{j'}$, which are smaller in this sense. But those $R'_{j'}$ are, in turn, generated from smaller R_i , and so, in this last alternative, R_1 is generated from strictly smaller R_i , which contradicts canonicity. Arguing symmetrically we see that the R_j and the $R'_{j'}$ coincide. \square

This theorem yields completeness for the set of transition relations formed from Spec expanded with an additional species name. As the additional species name is only used in the case that P and Q are open, the theorem holds generally for closed processes. The possibility of expanding Spec also ensures that the equivalence of (1) and (2) holds without any assumption. It is worth noting that it follows from the proof that one can restrict the environments in (3) to be definable, by which we mean that all their values are definable.

One can, of course, directly define a transition relation in TRel on every net in Net . Specifically to every net $N = (T, \text{pre}, \text{post})$ one assigns the following transition relation:

$$\mathcal{R}(N) =_{\text{def}} \{(Y + \sum_{t \in T} X(t) \cdot t, Y + \sum_{t \in T} X(t) t) \mid X \in \mathcal{M}_f(T)\} (= \{(\cdot t, t)\}^*)$$

The function $\mathcal{R} : \text{Net} \rightarrow \text{TRel}$ preserves all the semantical structure:

Proposition 4.3. *For any rule R we have:*

$$\text{TRel}\llbracket R \rrbracket = \mathcal{R}(\text{Net}\llbracket R \rrbracket)$$

and, further, \mathcal{R} is a monoid homomorphism.

Consequently, the Petri net and transition relation semantics are consistent with the usual multi-transition semantics of Petri nets in that, for any qualitative process P and any Net process environment ρ , we have:

$$\text{TRel}\llbracket P \rrbracket(\mathcal{R} \circ \rho) = \mathcal{R}(\text{Net}\llbracket P \rrbracket(\rho))$$

We can think of \mathcal{R} as giving the semantics of Petri nets. As syntactic objects, Petri nets are generated from transitions by finite composition, and \mathcal{R} preserves that structure. The qualitative calculus provides a syntax for nets with its rules and finite composition structure and the completeness of the axiom system identifies nets as equivalence classes of terms, and so as a kind of non-free syntax.

There are other possible transition relations one can associate to Petri nets, such as transitions occurring by virtue of a single transition firing, or by virtue of a set of transitions firing, or by virtue of a maximal set (or multiset) of transitions firing. Of these, the first two are *modular*, meaning that they can be equipped with a suitable monoidal structure, but the last two are not. We chose the multiset semantics as it seemed the most natural given that one is in any case working with markings as multisets. Another variant would have been to decorate the transition relation with the set or multiset of transitions involved, but we decided against that as we are keeping transitions anonymous.

Petri nets being at once graphs and supporting the fundamentals of chemistry themselves provide a fundamental graphical notation. Biologists make much use of informal graphical means of describing pathways; further, there has been work by Kohn, Kitano and others, see, e.g., [21, 20, 22], to produce more formal graphical notations adequate for flexible natural pathway description. By taking abbreviatory conventions seriously, meaning having them as additional language constructs, one can give a different graphical semantics than Petri nets, for example Petri nets extended with reversible transitions. In that way one's formalism would begin to make contact with the more informal graphical approaches, but still retain direct contact with the more basic Petri nets, and their various semantics.

4.3. Quantitative Semantics: ODE. We give a semantics of the quantitative calculus in terms of ordinary differential equations (ODEs). Here one associates to a process a set of ODEs of the form:

$$\begin{aligned} \frac{d[S_1]}{dt} &= p_1 \\ &\vdots \\ \frac{d[S_m]}{dt} &= p_m \end{aligned}$$

where, as usual, we write $[S]$ for the concentration of a species S , and S_1, \dots, S_m is an enumeration of the species in Spec , and p_1, \dots, p_m are real polynomials over the $[S_i]$. We need to make a commutative monoid Diff out of such sets of ODEs and to that end we identify them with maps from Spec to the commutative monoid of real polynomials over the $[S_i]$. They then have the standard pointwise monoid structure:

$$(\mathbf{p} \mid_{\text{Diff}} \mathbf{q})(S) = \mathbf{p}(S) + \mathbf{q}(S)$$

and:

$$\text{NIL}_{\text{Diff}}(S) = 0$$

It only remains to give the semantics of rules:

$$\text{Diff}\llbracket X \xrightarrow{r} Y \rrbracket(S) = r(Y(S) - X(S)) \prod_{S \in \text{Spec}} [S]^{X(S)}$$

Not all sets of ODEs of the above form can arise as the semantics of (closed) processes or, equivalently, using Proposition 4.6 below, of quantitative Petri nets. However there is a simple condition, characterising those that are so definable. Expand each p_i as a linear sum of distinct monomials: $\sum_j a_{ij} q_{ij}$ ($a_i \neq 0$). We then have the following theorem, due to Hárs and Tóth; we include a proof for the sake of completeness.

Theorem 4.4. *A set of ODEs of the above form is definable if, and only if, whenever any a_{ij} is negative then the power of $[S_i]$ in q_{ij} is non-zero.*

Proof. It is easy to see that this condition holds for all definable sets of equations. For the converse, it is enough to show that, for any i, j , we can define a rule R_{ij} such that $\text{Diff}\llbracket R_{ij} \rrbracket(S) = 0$, if $S \neq S_i$, and $\text{Diff}\llbracket R_{ij} \rrbracket(S) = a_{ij} q_{ij}$. With that, the parallel composition of the R_{ij} defines the given set of differential equations.

So, let q_{ij} have the form $\prod_{k=1,n} [S_k]^{m_k}$. There are two cases. Suppose first that a_{ij} is negative. Then $m_i \neq 0$ and we take the rule R_{ij} to be

$$m_1 S_1 + \cdots + m_n S_n \xrightarrow{a_{ij}/m_i} m_1 S_1 + \cdots + m_{i-1} S_{i-1} + m_{i+1} S_{i+1} + \cdots + m_n S_n$$

The other case is where a_{ij} is positive, when we take the rule R_{ij} to be

$$m_1 S_1 + \cdots + m_n S_n \xrightarrow{a_{ij}} S_i + (m_1 S_1 + \cdots + m_n S_n)$$

□

There can be non-trivial examples of different quantitative (closed) processes (or, equivalently systems of reactions, or, as we see below, non-isomorphic quantitative Petri nets) having the same ODE semantics. The example

$$2A \xrightarrow{2r} B + A = 2A \xrightarrow{r} 2B$$

was (essentially) given in [3], and another example was given in [12]; see [33, 34] for work on finding optimal reaction systems having given ODEs. We now axiomatise this equivalence. The following axioms hold:

$$\begin{aligned} X &\xrightarrow{r} X &= & \text{NIL} \\ X &\xrightarrow{r+s} Z &= & X \xrightarrow{r} Z \mid X \xrightarrow{s} Z \\ X + Y &\xrightarrow{r} X + Z &= & X + Y \xrightarrow{r} X \mid X + Y \xrightarrow{r} X + Y + Z \\ X + Y + Z &\xrightarrow{r} X &= & X + Y + Z \xrightarrow{r} X + Y \mid X + Y + Z \xrightarrow{r} X + Z \\ X &\xrightarrow{r} X + Y + Z &= & X \xrightarrow{r} X + Y \mid X \xrightarrow{r} X + Z \\ X + Y &\xrightarrow{r} X \mid X + Y \xrightarrow{s} (X + Y) + Y &= & \text{NIL} & (\text{if } r = s) \\ X + Y &\xrightarrow{r} X \mid X + Y \xrightarrow{s} (X + Y) + Y &= & X + Y \xrightarrow{r-s} X & (\text{if } r > s) \\ X + Y &\xrightarrow{r} X \mid X + Y \xrightarrow{s} (X + Y) + Y &= & X + Y \xrightarrow{s-r} (X + Y) + Y & (\text{if } r < s) \end{aligned}$$

The first two of these axioms are easy to understand. For the others, let us say that a rule of the form $X + Y \xrightarrow{r} X$ is of *consumption* type and that a rule of the

form $X \xrightarrow{r} X + Y$ is of *production* type. Then the third axiom divides an arbitrary rule into one of each type; the fourth axiom divides a consumption rule into two other consumption rules that are simpler in the sense that less is consumed in each; the fifth axiom divides a rule of production type into two simpler such rules; and the last three concern the balance between production and consumption when a consumption rule is put in parallel with a production one.

The six consumption and production axioms may be easier to understand if we write $Y \xrightarrow[r]{r} Z$ for $X + Y \xrightarrow{r} X + Z$. Consumption and production axioms then take the respective forms $Y \xrightarrow[r]{r} \emptyset$ and $\emptyset \xrightarrow[r]{r} Y$, and the six axioms become:

$$\begin{aligned}
Y \xrightarrow[r]{r} Z &= Y \xrightarrow[r]{r} \emptyset \mid \emptyset \xrightarrow[r]{r} Z \\
Y + Z \xrightarrow[r]{r} \emptyset &= Z \xrightarrow[r]{r} \emptyset \mid Y \xrightarrow[r]{r} \emptyset \\
\emptyset \xrightarrow[r]{r} Y + Z &= \emptyset \xrightarrow[r]{r} Y \mid \emptyset \xrightarrow[r]{r} Z \\
Y \xrightarrow[r]{r} \emptyset \mid \emptyset \xrightarrow[s]{s} Y &= \text{NIL} \quad (\text{if } r = s) \\
Y \xrightarrow[r]{r} \emptyset \mid \emptyset \xrightarrow[s]{s} Y &= Y \xrightarrow[r-s]{r-s} \emptyset \quad (\text{if } r > s) \\
Y \xrightarrow[r]{r} \emptyset \mid \emptyset \xrightarrow[s]{s} Y &= \emptyset \xrightarrow[s-r]{s-r} Y \quad (\text{if } r < s)
\end{aligned}$$

Note that in all the equations the left-hand-side of the rules do not change: in terms of the ODEs we are only reorganising the contributions to the coefficients of the monomials. There is a canonical form. Say that a consumption, respectively production, rule is *unary* if it is of the form $S \xrightarrow[r]{r} \emptyset$, respectively $\emptyset \xrightarrow[r]{r} S$. Then a process is in canonical form if it is a composition of process names and such unary rules, with no species having both a consumption and a production rule. Every process can be put in canonical form: one eliminates trivial rules using the first axiom, reduces rules to compositions of unary ones using the next three axioms, and combines unary rules with identical left- and right-hand sides using the next, and finally uses the last two rules to ensure that no species has both a consumption and a production rule.

With the addition of the above axioms, the logic of the quantitative calculus is complete for the ODE semantics:

Theorem 4.5. *The following are equivalent for any two processes P and Q :*

- (1) *P and Q can be proven equal using the above axioms for the ODE semantics.*
- (2) *Any two canonical forms of P and Q are the same up to reordering.*
- (3) *$\text{Diff}\llbracket P \rrbracket(\rho) = \text{Diff}\llbracket Q \rrbracket(\rho)$ for all ρ .*

Proof. As always, we need only assume (3) and prove (2). So choose canonical forms $A_1 \mid \dots \mid A_m \mid R_1 \mid \dots \mid R_n$ and $A'_1 \mid \dots \mid A'_{m'} \mid R'_1 \mid \dots \mid R'_{n'}$ of P and Q . A consumption

rule $X + S \xrightarrow{r} X$ has semantics with value 0 everywhere except at S where it has value

$$-r[S] \prod_i [S_i]^{X(S_i)}$$

A production rule $X \xrightarrow{r} X + S$ has semantics with value 0 everywhere except at S where it has value

$$r \prod_i [S_i]^{X(S_i)}$$

So, considering $\text{Diff}\llbracket P \rrbracket(\rho)$ and $\text{Diff}\llbracket Q \rrbracket(\rho)$, with ρ having value 0 everywhere, we see that the R_i and the R'_i must coincide.

To show that A_1 , say, occurs with the same multiplicity in the two canonical forms one employs the environment ρ , which has value 0 everywhere except at A_1 where it has value $\text{Diff}\llbracket nS \xrightarrow{1} \emptyset \rrbracket$ where S is a chosen element of Spec and $n > 0$ is chosen greater than the multiplicity of S in the left-hand side of any R_i . \square

As before, it follows from the proof that one can restrict the environments in condition (3) to be definable.

One can directly assign ODE semantics to Petri nets in CNet, as was already done more generally in [14]. We set:

$$\mathcal{D}(T, \text{pre}, \text{post}, \text{rate})(S) = \sum_{t \in T} \text{rate}(t)(t'(S) - t(S)) \prod_{S' \in \text{Spec}} [S']^{t(S')}$$

thereby defining the semantics $\mathcal{D} : \text{CNet} \rightarrow \text{Diff}$. As before this Petri net semantics preserves all the semantical structure:

Proposition 4.6. *For any rule R we have:*

$$\text{Diff}\llbracket R \rrbracket = \mathcal{D}(\text{CNet}\llbracket R \rrbracket)$$

Further, $\mathcal{D} : (\text{CNet}, |\text{CNet}|, \text{NIL}_{\text{CNet}}) \rightarrow (\text{Diff}, |\text{Diff}|, \text{NIL}_{\text{Diff}})$ is a monoid homomorphism.

So the Petri net and ODE semantics are consistent in that, for any quantitative process P and any CNet process environment ρ , we have:

$$\text{Diff}\llbracket P \rrbracket(\mathcal{D} \circ \rho) = \mathcal{D}(\text{CNet}\llbracket P \rrbracket(\rho))$$

We now turn to our full abstraction results for the ODE semantics. We begin with a notion of “same simulation behaviour.” For any $\mathbf{p}, \mathbf{q} \in \text{Diff}$, define $\mathbf{p} \sim \mathbf{q}$ to hold if, and only if, for all non-negative initial values (i.e., those in $\mathbf{x} \in \mathbb{R}_{\geq 0}^{|\text{Spec}|}$) there is a time interval $[0, t)$ on which the ODEs given by \mathbf{p} and \mathbf{q} have the same solutions with initial value \mathbf{x} . Thus $\mathbf{p} \sim \mathbf{q}$ if the ODEs given by \mathbf{p} and \mathbf{q} have the same solutions locally, for any given initial value. This relation is evidently an equivalence; it seems a reasonable definition as, in general, ODEs need not have global solutions.

One concern about the definition of same simulation behaviour is that solutions of the ODEs given by \mathbf{p} may have negative components, even although the initial value is non-negative. However this does not happen for definable \mathbf{p} , as, in that case, solutions remain non-negative, as shown in [35, 16]; one can prove this using an immediate consequence of the above definability condition, that if $[S_i] = 0$, then $\mathbf{p}_i \geq 0$.

Next we define the set of process contexts C, D by the following abstract syntax:

$$C ::= R \mid C|D \mid \text{NIL} \mid A = C; D \mid A \mid [\cdot]$$

which is the same as that for processes except that the possibility of a “hole” $[\cdot]$ has been added. Given any context C and process P one obtains a process $C[P]$ by replacing all the holes in C by P . We then define our notion of contextual equivalence by:

$$P \approx Q \text{ iff } \forall C. (C[P] \text{ and } C[Q] \text{ closed} \Rightarrow \text{Diff}[C[P]] \sim \text{Diff}[C[Q]])$$

We now show that our semantics is indeed fully abstract, in that contextual equivalence coincides with having the same ODE semantics. First we need a lemma.

Lemma 4.7. *For any $\mathbf{p}, \mathbf{q} \in \text{Diff}$, $\mathbf{p} \sim \mathbf{q}$ holds if, and only if, $\mathbf{p} = \mathbf{q}$.*

Proof. Assume $\mathbf{p} \sim \mathbf{q}$. Fix an initial value \mathbf{x} . By a standard existence theorem for ordinary differential equations (either that of Cauchy-Peano or that of Picard-Lindelöf), the ODEs given by \mathbf{p} have a solution on some $[0, t)$ with the given initial value, and so, by the assumption, the ODEs given by \mathbf{q} have the same solution. Such solutions determine the values of the $\frac{d[S_i]}{dt}$ at the given initial value, and so the polynomials $\mathbf{p}([S_i])$ and $\mathbf{q}([S_i])$ have the same values at the given initial value. As the initial value was chosen arbitrarily from $\mathbb{R}_{\geq 0}^{|\text{Spec}|}$, it follows that the two polynomials are identical. \square

Theorem 4.8. *For any processes P and Q we have:*

$$P \approx Q \quad \text{iff} \quad \forall \rho. \text{Diff}[P](\rho) = \text{Diff}[Q](\rho)$$

Proof. For the implication from right to left, assume that $\forall \rho. \text{Diff}[P](\rho) = \text{Diff}[Q](\rho)$. It is then easy to show for any context C that $\forall \rho. \text{Diff}[C[P]](\rho) = \text{Diff}[C[Q]](\rho)$; the proof is by induction on the structure of C .

Conversely, suppose that $P \approx Q$. Let ρ be a definable environment, and suppose that $\rho(A) = \text{Diff}[P_A]$; and let A_1, \dots, A_n be a list without repetition of all the free process variables of P or Q . Define C be the context

$$\begin{aligned} A_1 &= P_{A_1}; \\ &\vdots \\ A_n &= P_{A_n}; \\ [\cdot] \end{aligned}$$

Note that both $C[P]$ and $C[Q]$ are closed; this lets us apply the assumption:

$$\begin{aligned} \text{Diff}[P](\rho) &= \text{Diff}[P](\rho[A_1 \mapsto \text{Diff}[P_{A_1}]] \dots [A_n \mapsto \text{Diff}[P_{A_n}]]) \\ &= \text{Diff}[P[P_{A_n}/A_n] \dots [P_{A_1}/A_1]](\rho) && \text{(substitution and denotation commute)} \\ &= \text{Diff}[C[P]](\rho) && \text{(using the proof rules)} \\ &= \text{Diff}[C[Q]](\rho) && \text{(by assumption and Lemma 4.7)} \\ &= \text{Diff}[Q](\rho) \end{aligned}$$

The conclusion then follows by Theorem 4.5, and the remark after it on definable environments. \square

4.4. Quantitative Semantics: Stochastic. We give another semantics of the quantitative calculus in terms of the set SMatrix of *stochastic transition matrices*. These are maps $Q : \mathcal{M}_f(\text{Spec})^2 \rightarrow \mathbb{R}$ such that

$$Q(X, Y) \geq 0 \quad (\text{if } Y \neq X)$$

and

$$Q(X, X) = - \sum_{Y \neq X} Q(X, Y)$$

This has the following pointwise commutative monoid structure:

$$(Q \mid_{\text{SMatrix}} Q')(X, Y) = Q(X, Y) + Q'(X, Y)$$

$$\text{NIL}_{\text{SMatrix}}(X, Y) = 0$$

For the semantics of rules, first define $(X, Y) \preceq (X', Y')$ to hold for multisets X, Y, X', Y' iff, for some (necessarily unique) Z we have $X' = Z + X$ and $Y' = Z + Y$. Then, in case $Y \neq X$ and $X' \geq X$, we put:

$$\text{SMatrix}[X \xrightarrow{r} Y](X', Y') = \begin{cases} r \binom{X'}{X} & (\text{if } (X, Y) \preceq (X', Y')) \\ -r \binom{X'}{X} & (\text{if } Y' = X') \\ 0 & (\text{otherwise}) \end{cases}$$

where

$$\binom{X'}{X} =_{\text{def}} \prod_{S \in \text{Spec}} \binom{X'(S)}{X(S)}$$

In all other cases we put:

$$\text{SMatrix}[X \xrightarrow{r} Y](X', Y') = 0$$

We do not know any characterisation of the stochastic transition matrices definable by nets. One necessary condition is that there is a k such that for all X we have $|\{Y \mid Q(X, Y) > 0\}| \leq k$. However a first problem seems to be, allowing nets with countably many transitions, to characterise the stochastic transition relations defined by such nets.

The following axioms hold:

$$X \xrightarrow{r} X = \text{NIL}$$

$$X \xrightarrow{r+s} Z = X \xrightarrow{r} Z \mid X \xrightarrow{s} Z$$

There is an evident canonical form: a composition of process names and rules in which no rule of the form $X \xrightarrow{r} X$ occurs and at most one rule with any given left- and right-hand side occurs. With the addition of the above two axioms, the logic of the quantitative calculus is complete for the stochastic semantics:

Theorem 4.9. *The following are equivalent for any two processes P and Q :*

- (1) P and Q can be proven equal using the above axioms for the stochastic semantics.
- (2) Any two canonical forms of P and Q are the same up to reordering.
- (3) $\text{SMatrix}[P](\rho) = \text{SMatrix}[Q](\rho)$ for all ρ .

Proof. As always, we need only assume (3) and prove (2). So choose canonical forms $A_1 \mid \dots \mid A_m \mid R_1 \mid \dots \mid R_n$ and $A'_1 \mid \dots \mid A'_{m'} \mid R'_1 \mid \dots \mid R'_{n'}$ of P and Q . Setting ρ to be constantly $\text{NIL}_{\text{SMatrix}}$ for any $X, Y \in \mathcal{M}_f(\text{Spec})$ with $X \neq Y$, we have:

$$\text{SMatrix}[P](\rho)(X, Y) = \sum_{(X_j, Y_j) \preceq (X, Y)} r_j \binom{X}{X_j}$$

where $R_j = X_j \xrightarrow{r_j} Y_j$; a similar formula holds for Q , with $R'_k = X'_k \xrightarrow{r'_k} Y'_k$.

We prove that $X \xrightarrow{r} Y$ occurs in one canonical form iff it is in the other by induction on $\|X\|$. So consider an $R_j = X_j \xrightarrow{r_j} Y_j$, to show it is some R'_k . From the above formula for $\text{SMatrix}[P]$ we have:

$$\text{SMatrix}[P](\rho)(X_j, Y_j) = \sum_{\substack{(X_{j'}, Y_{j'}) \preceq (X_j, Y_j) \\ \|X_{j'}\| < \|X_j\|}} r_{j'} \binom{X}{X_{j'}} + r_j \binom{X_j}{X_j}$$

using the fact that if $(X', Y') \preceq (X, Y)$ and $\|X'\| = \|X\|$ then $X' = X$ and $Y' = Y$. Similarly, we have:

$$\text{SMatrix}[Q](\rho)(X_j, Y_j) = \sum_{\substack{(X'_k, Y'_k) \preceq (X_j, Y_j) \\ \|X'_k\| < \|X_j\|}} r'_k \binom{X}{X'_k} + \delta$$

where $\delta = 0$ unless there is a (necessarily unique) $R'_k = X'_k \xrightarrow{r'_k} Y'_k$ with $X'_k = X_j$ and $Y'_k = Y_j$, when $\delta = r'_k \binom{X'_k}{X'_k} (= r'_k \binom{X_j}{X_j})$.

We know that $\text{SMatrix}[P](\rho)(X_j, Y_j) = \text{SMatrix}[Q](\rho)(X_j, Y_j)$, so, applying the induction hypothesis, we see that $r_j \binom{X_j}{X_j} = \delta$. So, as $r_j \neq 0$, we have $r_j \binom{X_j}{X_j} = \delta = r'_k \binom{X_j}{X_j}$, and so $r'_k = r_j$. Therefore R_j occurs in the canonical form of Q , as required. The converse assertion, that rules appearing in the canonical form of Q also appear in the canonical form of P is proved similarly.

To see that the same process identifiers occur in both canonical forms, and with the same multiplicities, suppose that A has multiplicity m in P and multiplicity m' in Q . Define ρ to have value $\text{NIL}_{\text{SMatrix}}$ everywhere except at A where it has value $\text{SMatrix}[nS \xrightarrow{1} \emptyset]$, for chosen $S \in \text{Spec}$, and $n > 0$ greater than the multiplicity of S in the left-hand side of any R_j or $R_{j'}$. Noting that $(X', Y') \preceq (X, \emptyset)$ iff $X' = X$ and $Y' = \emptyset$ we see that $\text{SMatrix}[R_1 \mid \dots \mid R_n](\rho)(nS, \emptyset) = 0$ and so $\text{SMatrix}[P](\rho)(nS, \emptyset) = mn!$. Arguing similarly, we see that $\text{SMatrix}[Q](\rho)(nS, \emptyset) = m'n!$ where m' is the multiplicity of A_1 in the canonical form of Q . It follows from the assumption that $m' = m$. \square

Again as before, it follows from the proof that one can restrict the environments in condition (3) to be definable.

There is a standard Petri net stochastic semantics $\mathcal{S} : \text{CNet} \rightarrow \text{SMatrix}$, see, e.g., [36, 14]. It is given for $Y \neq X$ by:

$$\mathcal{S}(T, \text{pre}, \text{post}, \text{rate})(X, Y) = \sum_{t \in T} \{ \text{rate}(t) \binom{X}{\cdot_t} \mid (\cdot_t, t) \preceq (X, Y) \}$$

and thereby determined on the diagonal.

Proposition 4.10. *For any rule R we have:*

$$\text{SMatrix}[\![R]\!] = \mathcal{S}(\text{CNet}[\![R]\!])$$

and \mathcal{S} is a monoid homomorphism.

It follows that the Petri net and stochastic semantics are consistent in that, for any quantitative process P and any CNet process environment ρ , we have:

$$\text{SMatrix}[\![P]\!](\mathcal{S} \circ \rho) = \mathcal{S}(\text{CNet}[\![P]\!](\rho))$$

As in the case of the ODE semantics, we can consider notions of contextual equivalence for stochastic simulation. The analogue of a solution of the ODEs with a given initial value is the Markov process induced by a stochastic matrix $Q \in \text{SMatrix}$, with a given initial point-mass probability distribution δ_X . (For an explanation of how Markov processes arise from stochastic matrices see, for example, Chapter 2 of [27].)

For any $Q, Q' \in \text{SMatrix}$, define $Q \sim Q'$ to hold if, and only if, for any given initial point-mass probability distribution δ_X , Q and Q' induce the same Markov process. Contextual equivalence is then defined by:

$$P \approx Q \text{ iff } \forall C. (C[P] \text{ and } C[Q] \text{ closed} \Rightarrow \text{SMatrix}[\![C[P]]\!] \sim \text{SMatrix}[\![C[Q]]\!])$$

Lemma 4.11. *For any $Q, Q' \in \text{SMatrix}$, $Q \sim Q'$ holds if, and only if, $Q = Q'$.*

Proof. The result is an immediate consequence of the fact that the X -th row of any $Q \in \text{SMatrix}$ is determined by the induced Markov process starting at δ_X . To see this, first note that the first holding time of this process is exponentially distributed with parameter $Q_X =_{\text{def}} -Q(X, X)$. If Q_X is 0 then so is the entire X -th row of Q . Otherwise, as the jump chain of the Markov process is the X -th row of the jump matrix Π of Q , and as $\Pi_{X,Y} = Q(X, Y)/Q_X$ off the diagonal when $Q_X \neq 0$, the X -th row of Q is again determined. \square

Using this lemma, we have that the stochastic matrix semantics is indeed fully abstract; the proof is entirely analogous to that of Theorem 4.8.

Theorem 4.12. *For any processes P and Q we have:*

$$P \approx Q \text{ iff } \forall \rho. \text{SMatrix}[\![P]\!](\rho) = \text{SMatrix}[\![Q]\!](\rho)$$

ACKNOWLEDGEMENTS

We would like to thank Michael Pedersen and János Tóth for helpful discussions.

REFERENCES

- [1] Albert-László Barabási, and Zoltán N. Oltvai, Network biology: understanding the cell's functional organization, *Nature Reviews Genetics*, **5**(2):101–13, 2004.
- [2] Laurence Calzone, François Fages, and Sylvain Soliman, BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge, *Bioinformatics* **22**(14), 1805–1807, 2006.
- [3] Luca Cardelli, On process rate semantics, *Theor. Comput. Sci.*, **391**(3), 190–215, 2008.
- [4] Claudine Chaouiya, Petri net modelling of biological networks, *Briefings in Bioinformatics*, **8**(4), 210–219, 2007.

- [5] Gheorghe Craciun and Martin Feinberg, Multiple equilibria in complex chemical reaction networks: II. The species-reaction graph, *SIAM J. Appl. Math.*, **66**(4), 1321–1338, 2006.
- [6] Vincent Danos and Cosimo Laneve, Formal molecular biology, *Theor. Comput. Sci.* **325**(1), 69–110, 2004.
- [7] Péter Érdi and János Tóth, *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models*, Princeton University Press, 1989.
- [8] Mirela Domijan and Markus Kirkilionis, Graph theory and qualitative analysis of reaction networks, *Networks and heterogeneous media*, **3**(2), 295–322, American Institute of Mathematical Sciences, 2008.
- [9] Marcelo P. Fiore, Achim Jung, Eugenio Moggi, Peter O’Hearn, Jon Riecke, Giuseppe Rosolini, and Ian Stark, Domains and denotational semantics: history, accomplishments and open problems, *Bulletin of the European Association for Theoretical Computer Science*, **59**, 227–256, 1996.
- [10] Timothy P. Garrington and Gary L. Johnson, Organization and regulation of mitogen-activated protein kinase signaling pathways, *Current Opinion in Cell Biology*, **11**, 211–218, 1999.
- [11] Jan Friso Groote and Marc Voorhoeve, Operational semantics for Petri net components, *Theor. Comput. Sci.*, **379**(1–2), 1–19, 2007.
- [12] Vera Hárs and János Tóth, On the inverse problem of reaction kinetics, *Qualitative Theory of Differential Equations* (eds. M. Farkas and L. Hatvani) Coll. Math. Soc. J. Bolyai, **30**, 363–379, North-Holland, 1981.
- [13] Leland H. Hartwell, John J. Hopfield, Stanislas Leibler, and Andrew W. Murray, From molecular to modular cell biology. *Nature* **402**, C47–C52, 1999.
- [14] Monika Heiner, David Gilbert, and Robin Donaldson, Petri nets for systems and synthetic biology, *Proc. SFM 2008* (eds. Marco Bernardo, Pierpaolo Degano, and Gianluigi Zavattaro), LNCS **5016**, 215–264, Springer, 2008.
- [15] William S. Hlavacek, James R. Faeder, Michael L. Blinov, Richard G. Posner, Michael Hucka, and Walter Fontana, Rules for modeling signal-transduction systems, *Sci. STKE*, **2006**(344), p. re6, 2006.
- [16] Sergei I. Hudjaev and Aizik I. Vol’pert, *Analysis in classes of discontinuous functions and equations of mathematical physics*, Mechanics: analysis, **8**, Dordrecht, 1985.
- [17] A. N. Ivanova, Conditions for uniqueness of stationary state of kinetic systems related to structural scheme of reactions, *Kinet. Katal.*, **20**(4), 1019–1023, 1979.
- [18] Hans-Michael Kaltenbach and Jörg Stelling, Modular analysis of biological networks, *Advances in Systems Biology*, Advances in Experimental Medicine and Biology, **736**, Part 1, 3–17, Springer, 2012.
- [19] Kurt Jensen, *Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use*, Volume 1, Monographs in Theoretical Computer Science, Springer, 1992.
- [20] Hiroaki Kitano, A graphical notation for biochemical networks, *BIOSILICO*, **1**(5), 169–176, 2003.
- [21] Kurt W. Kohn, Molecular interaction map of the mammalian cell cycle control and DNA repair systems, *Molecular Biology of the Cell*, **10**, 2703–2734, 1999.
- [22] Kurt W. Kohn, Mirit I. Aladjem, John N. Weinstein, and Yves Pommier, Molecular interaction maps of bioregulatory networks: a general rubric for systems biology, *Mol. Biol. Cell* **17**, 1–13, 2006.
- [23] Robin Milner, *Communication and Concurrency*, Prentice Hall, 1989.
- [24] Robin Milner, *Communicating and Mobile Systems - The Pi-Calculus*, CUP, 1999.
- [25] Hiroshi Matsuno, Chen Li, and Satoru Miyano, Petri net based descriptions for systematic understanding of biological pathways, *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, **89-A**(11), 3166–3174, 2006.
- [26] Eric Mjolsness and Guy Yosiphon, Stochastic process semantics for dynamical grammars, *Ann. Math. Artif. Intell.*, **47**(3–4), 329–395, 2006.
- [27] James R. Norris, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics, **2**, CUP, 1998.
- [28] Michael Pedersen, Compositional denitions of minimal flows in Petri nets, *Proc. 6th CMSB* (eds. Monika Heiner and Adelinde M. Uhrmacher), LNCS **5307**, 288–307, Springer, 2008.

- [29] Michael Pedersen and Gordon D. Plotkin, A language for biochemical systems: design and formal specification. *T. Comp. Sys. Biology*, **12**, 77–145, 2010.
- [30] Corrado Priami and Paola Quaglia, Beta binders for biological interactions, *Proc. 2nd. CMSB* (eds. Vincent Danos and Vincent Schächter), LNCS **3082** 20–33, Springer, 2005.
- [31] Corrado Priami, Aviv Regev, Ehud Y. Shapiro, and William Silverman, Application of a stochastic name-passing calculus to representation and simulation of molecular processes. *Inf. Process. Lett.* **80**(1), 25–31, 2001.
- [32] Aviv Regev, Ekaterina M. Panina, William Silverman, Luca Cardelli, and Ehud Y. Shapiro, BioAmbients: an abstraction for biological compartments, *Theor. Comput. Sci.* **325**(1), 141–167, 2004.
- [33] Gabor Szederényi, Computing sparse and dense realizations of reaction kinetic systems, *Journal of Mathematical Chemistry*, **47**, 551–568, 2009.
- [34] Gabor Szederényi, Katalin M. Hangos, and Tamas Péni, Maximal and minimal realizations of reaction kinetic systems: Computation and properties, *MATCH Communications in mathematical and in computer chemistry*, **65**(2), 2011, also available as arXiv:1005.2913v1 [q-bio.MN].
- [35] Aizik I. Vol’pert, Differential equations on graphs, *Mathematics of the USSR-Sbornik*, **17**(4), 571–582, 1972.
- [36] Darren J. Wilkinson, *Stochastic Modelling for System Biology*, CRC Press, New York, 2006.
- [37] Glynn Winskel, *The Formal Semantics of Programming Languages*, MIT Press, 1993.
- [38] Chin-Rang Yang, Bruce E. Shapiro, Eric Mjolsness, and G. Wesley Hatfield, An enzyme mechanism language for the mathematical modeling of metabolic pathways, *Bioinformatics* **21**(6), 774–780, 2005.